

DOCUMENT RESUME

ED 293 252

EC 202 451

AUTHOR Walsh, Patricia C.; And Others
TITLE Cluster Score Performance of Learning Disabled Students on the Woodcock-Johnson Psycho-Educational Battery.
PUB DATE Apr 87
NOTE 20p.; Paper presented at the Annual Convention of the Council for Exceptional Children (65th, Chicago, IL, April 20-24, 1987).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Evaluative/Feasibility (142)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Ability; *Academic Achievement; Cognitive Processes; Elementary Education; *Handicap Identification; *Learning Disabilities; *Test Interpretation; *Test Norms
IDENTIFIERS Woodcock Johnson Psycho Educational Battery

ABSTRACT

The study evaluated the Woodcock-Johnson Psycho-Educational Battery (WJPEB) with 71 elementary grade learning disabled (LD) students. The following questions were addressed: (1) Are there significant discrepancies between ability and achievement, as measured by the WJPEB, for LD children, and if so, in what academic areas? (2) Are there patterns of strengths and deficits on the Aptitude and Ability Cluster scores of LD students? (3) Do the WJPEB Broad Cognitive Ability scores predict achievement as well as the individual Aptitude Cluster scores? and (4) How does the performance of previously identified learning disabled students compare to the WJPEB normative data?. Results indicated: (1) limited effectiveness in identifying LD students (only 60-68% were identified by the test); (2) a pattern of deficits in the areas of memory, written language aptitude, and knowledge aptitude; (3) inability of the Broad Cognitive Ability Score to identify certain problem areas (e.g. reading deficit); and (4) average performance of the subjects on tests of cognitive abilities and below national norm performance in achievement areas. It is concluded the WJPEB was neither efficient nor effective in identifying this sample of learning disabled students, and that standard scoring procedures, rather than the Broad Cognitive Ability score, should be used. (DB)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED293252

☒ This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official
DERI position or policy.

Cluster Score Performance
of Learning Disabled Students on the Woodcock-Johnson
Psycho-Educational Battery

Paper Presented at the 65th Annual Council for Exceptional
Children Convention, Chicago, Illinois, April, 1987

Patricia C. Walsh, Chairperson and Professor
Barbara Lowenthal, Associate Professor
Glen Thompson, Professor
Special Education Department
Northeastern Illinois University
5500 N. St. Louis Ave.
Chicago, Ill. 60625

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Glen R. Thompson

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

EC 202451

Cluster Score Performance
of Learning Disabled Students on the Woodcock-Johnson
Psycho-Educational Battery

INTRODUCTION AND RELATED RESEARCH

The Woodcock-Johnson Psycho-Educational Battery (WJPEB) is currently being utilized to assist in the identification of children and adults with learning disabilities (Compton, 1984). According to Woodcock (1978), the WJPEB is a set of wide age range, comprehensive, and individually administered standardized tests which measure cognitive abilities and scholastic aptitudes (WJTCA), achievement (WJTA), and interests.

One of the suggested assets of the WJPEB is its capacity to compare a student's aptitude to his/her current levels of achievement in four academic areas (Woodcock 1978; Hessler, 1984) and to determine if there are educationally significant discrepancies. This determination of aptitude-achievement discrepancies is essential to the process used in identifying students with learning disabilities (LD) according to definitions established under Public Law 94-142 (USOE, 1977). The validity of the WJTCA in the measurement of aptitude is of question in several studies. Results of some research (Coleman and Harmer, 1985; Bracken, Prasse, and Breen, 1984; Thompson and Brassard, 1984a, 1984b; Shinn, Algozzine, Marston, and Ysseldyke, 1982; Reeve, Hall, and Zakreski, 1979) suggest that LD students perform lower on the WJTCA Broad Cognitive Score than they do on the Wechsler Intelligence Scale For Children-Revised (WISC-R) (Wechsler, 1974). The range of mean difference scores was from 13

points (Reeve, et. al 1979) to 5.65 (Coleman and Harmer, 1985). Others have reported little mean difference in the WJTCa and WISC-R scores (Woodcock, 1984a). Shinn, et al, (1982) suggested that the WJTCa Broad Cognitive Score might actually measure achievement rather than ability and if so, this could well work to the disadvantage of LD students and result in confusion in the process of making placement decisions.

A few studies have reviewed the performance of LD students as compared to the WJPEB normative data. Shinn et al. (1982) and Bracken, Prasse, and Breen (1984) reported in their research that LD students scored significantly lower on the WJPEB than did the normative samples. These findings comparing WJTCa vs WISC-R and LD performance vs normative data raise serious concerns about the ability of the WJPEB to accurately depict LD students aptitude. As to the validity of the achievement clusters, Hall, Reeve, and Zakreski (1984) reported from their investigation that the WJPEB Tests of Achievement appear to be technically adequate in terms of concurrent validity for samples of elementary-aged LD students. The research findings suggest that the WJPEB aptitude clusters may have achievement loaded content and are therefore less able to identify potential in learning disabled students. This would greatly affect the ability of the WJPEB to establish aptitude-achievement discrepancies useful in the decision-making process.

Woodcock (1984b) pointed out that the WJTCa was also developed to provide information regarding special intellectual abilities that may influence school learning. Whether LD students do exhibit unique patterns

of strengths and deficits on the Woodcock Aptitude and Ability Clusters has been investigated. Results of some research (McGue, Shinn, and Ysseldyke, 1982; Reeve, Hall, and Zakreski, 1979) indicated that there was an apparent overlap in the four Aptitude Clusters. Correlations between clusters are high and the Antonyms-Synonyms Subtest appears in each cluster. Findings from another study (Shinn, Algozzine, Marston, and Ysseldyke, 1982) showed that LD children performed significantly lower than the normative group in the following subtests which comprise the Knowledge Aptitude Cluster: Memory for Sentences, Quantitative Concepts, Antonyms-Synonyms, and Analogies. Breen (1986) suggested from the results of his research that the Memory for Sentences subtest could be a major contributor to the cognitive skill deficits characteristic of LD children. Coleman and Harmer (1985) found that their sample of referral cases had significantly lower mean scores on the Perceptual Speed Cluster. Subsequent investigation by Breen (1986) suggested that the Perceptual Speed and Memory Cluster Scores may be sensitive to differences. Research studies have demonstrated a tendency on the part of LD students to have lower scores in two of the Ability areas on the WJTC: Perceptual-Speed and Memory.

The use of the Broad Cognitive Score, rather than the Aptitude Cluster scores, as a predictor of achievement in the academic areas of reading, mathematics, written language, and knowledge has been suggested. Cummings and Moscato (1984a, 1984b) in their review of the research concluded that the Broad Cognitive Score predicted the Achievement Clusters as well as the appropriate Aptitude clusters did and that the clinician can safely avoid

generating these additional scores. Estabrook (1983), in his test review of the WJPEB, concurred that the Broad Cognitive Score predicted achievement as well as the Aptitude Clusters did. A later study (Phelps and Rosso, 1985) also raised the question of the redundancy of these aptitude clusters as their findings indicated that every WJPEB Achievement Cluster correlated as well or better with the Broad Cognitive Ability score as with its corresponding Aptitude measure. Woodcock (1984b) refuted these claims stressing the fact that discrepancies not based on norms must be corrected for regression error.

The present research has attempted to shed more light on the issues raised here. Specifically:

1. Are there significant discrepancies between ability and achievement, as measured by the WJPEB, for LD children? If so, in what academic areas do these discrepancies lie?
2. Are there patterns of strengths and deficits on the Aptitude and Ability Clusters scores of LD students?
3. Do the WJPEB Broad Cognitive Ability scores predict achievement as well as the individual Aptitude Cluster scores?
4. How does the performance of previously identified learning disabled students compare to the WJPEB normative data?

METHOD

Subjects

Seventy-one learning disabled students, from ten elementary schools, in a school district located in a Northwest suburb of Chicago, participated

in this study. The students ranged in age from five years, eight months to fourteen years, three months, with an average age of eight years, 8-11 (S.D. = 27.08 months). Their grade range was from 1.2 to 8.5 with an average grade of 3.5 (S.D. = 2.05). The schools served middle class families whose ethnic backgrounds were the following: non-Hispanic White (80.6%), Black (.9%), American Indian (.1%), Asian-South Pacific (7.4%), and Hispanic (11.0%). The subjects received LD resource assistance and were identified as being learning disabled by an interdisciplinary team in accordance with district regulations and the mandate of P.L. 94-142 (Education of All Handicapped Children Act). Of the 71 subjects, 54 were boys (76%) and 17 were girls (24%).

Procedure

The WJPEB was administered to 71 subjects in resource rooms at their respective elementary schools by LD certified specialists who were employed by the school district. The data were analyzed using Statistics With Finesse (Bolding, 1981) with .05 level established as the point of statistical significance. In addition, case by case analysis was done to gain additional information regarding the WJPEB's performance with this sample of LD students. Specifically, two separate statistical counts were run, one using Aptitude standard scores and the other Broad Cognitive standard scores to determine aptitude-achievement discrepancies. These statistical manipulations used procedures outlined by Woodcock in DLM's Assessment Service Bulletin #5 (Woodcock, 1986) and as suggested by Woodcock (1984b) for those wishing to use the Broad Cognitive Score as the

predictor of discrepancies. Difference probabilities ranging between .01 and .16 were called moderate, and probabilities equal to or less than .01 were called severe for the DLM analysis.

RESULTS

Significant Discrepancies Between Aptitude and Achievement

Table I summarizes the results of the case by case analyses done on a sample of 71 students. The results from the procedure for finding discrepancies, described in the WJPEB Manual, were compared with the results from the two statistical methods described in DLM Service Bulletin # 5. The manual method identified 48 students with moderate to severe deficits, 31 were identified as having only one deficit, 15 with two, two with three, and none with four. Of the 31 students with reading deficits, 18 were moderate and 13 were severe. Similarly, in each area - Math, Written Language and Knowledge - there were fewer students with severe deficits. The statistical analysis using Aptitude standard scores to identify discrepancies indicated 47 cases with at least moderate deficits, 34 with only one deficit, 13 with two, none with three, and none with four. Thirty-one with deficits in reading were identified. The statistical count using Broad Cognitive standard scores, on the other hand, identified 41 cases with moderate to severe deficits, 32 with only one deficit, 7 with two deficits, one with three deficits, and one with four. Nineteen had deficits in reading. It is apparent that differences exist in the capacity of these three methods to identify discrepancies.

Reliabilities for difference scores were computed, and these range from .704 for Knowledge Aptitude and Knowledge to .905 for both Broad Cognitive Ability and Reading and Broad Cognitive Ability and Written Language. Four of these eight reliabilities (Math Aptitude and Math, Knowledge Aptitude and Knowledge, Broad Cognitive Ability and Math, and Broad Cognitive Ability and Knowledge) failed to reach the criterion level of .80 as suggested by Green (1981). The ability of the WJPEB to show discrepancies in student performance is inconsistent.

Patterns of Strengths and Deficits in the Ability and Aptitude Cluster Scores

An analysis of variance with repeated measures was performed using standard scores for Aptitudes and Abilities. The result of this analysis, as shown in Table II, was that the F-ratio for aptitude standard scores was significant. The Newman-Keuls procedure for making multiple comparisons was used to pinpoint the significant differences. The results of this procedure show FSBCA, Knowledge Aptitude, Memory, and Written Language Aptitude scores all to be significantly lower ($p < .01$) than Visual-Perceptual Speed (VPS), Math Aptitude, Reading Aptitude, Verbal Ability, and Reasoning. VPS and Math Aptitude scores were both lower than Verbal Ability and Reasoning.

A Comparison of Broad Cognitive Ability and Aptitude Scores in Predicting Achievement

Correlations of the standard scores for Broad Cognitive Ability and Aptitudes with Achievement are presented in Table III. Significant

($p < .01$) correlations were revealed. A comparison of Broad Cognitive Ability scores with Reading Aptitude scores in predicting Reading, with Math Aptitude scores in predicting Math, with Written Language scores in predicting Written Language, and with Knowledge Aptitude scores in predicting Knowledge was conducted. A statistical comparison of each of these four sets of correlations reveals no significant differences between the sets of correlations.

Normative Comparisons

Each of thirteen standard score means was compared with 100 using a one-sample Z test. The results of this analysis are summarized in Table IV and show that the means for the achievement areas of Written Language Achievement, Reading Achievement and Math Achievement were all significantly ($p < .01$) less than 100, while Knowledge Achievement was at norm. The means for Knowledge Aptitude was significantly below 100 ($p < .01$) and the Memory and Written Language Aptitude means were lower as well ($p < .05$).

Discussion and Conclusions

Significant Discrepancies Between Ability and Achievement

The case by case analysis revealed little significant difference in the identification of L.D. students whether one used the procedure suggested in the WJPEB Manual or the DLM statistical comparison of standard scores. The use of the Broad Cognitive Ability standard scores was more likely to under identify students especially in the area of reading, where only 5 moderate reading deficits were identified using the Broad Cognitive

standard score as the predictor. Using the DLM method with the Aptitude standard scores produced curious results in the area of written language. No students were identified as having a written language deficit given this procedure as compared to 12 moderate deficits with the Manual procedure. The inconsistencies that arose are difficult to explain at the present time. Possibly the probabilities used for the standard score comparisons need to be broadened.

In all three comparisons, the WJPEB identified from 60 - 68% of the students as having a learning disability. This limited effectiveness in identification suggests that practitioners refrain from using the WJPEB exclusively. The Aptitude subtests may well be measuring achievement and therefore excluding a third of the L.D. students in need of assistance. The process used in identification of the subjects in this study may also be affecting the results, underscoring a need for a specific, standard definition of learning disabilities that can be adhered to across school districts.

The only cognitive aptitude, or ability areas to show a deficit pattern were Memory, Written Language Aptitude, and Knowledge Aptitude. Memory mean scores, specifically, fell below Reasoning, Reading Aptitude, Math Aptitude, and Visual Perceptual Speed means. This sample also fell below standardization norms in the area of Memory. Since the Written Language and the Knowledge Aptitude Clusters include tests of short-term memory, this finding is consistent. Difficulty with short-term memory tasks is typical of LD students and seems readily identified by the WJPEB.

Again, the problem of uncertainty as to exactly how this sample of LD students was identified is an issue. If a deficit in memory was used as one criterion in identification, then the sample would show such a deficit. The most likely interpretation is that short-term memory problems are commonly found with students with learning disabilities whether Memory was a specific diagnostic category or not. Considering that Memory is the only deficit pattern found, diagnosticians need to evaluate whether or not the time taken administering the WJPEB is well spent.

These results are consistent with previous findings. Cluster subtests overlap and create a possibly misleading picture. The Written Language and Knowledge Clusters include subtests that require strong short-term memory skills. The effect of this is to present an overall picture of students with learning disabilities who have memory, written language and knowledge aptitude deficits when in fact they may be students with attention deficits and not specific problems with writing or knowledge.

Comparison of Full Scale and Aptitude Scores in Predicting Achievement

It seems from these results that there is little difference between Broad Cognitive Ability Scores and Aptitude scores in predicting achievement and in identifying students with LD. A comparison of correlations showed no significant difference between Broad Cognitive Ability scores and Aptitude scores in their correlations with achievement. The case by case analysis however suggests that the Broad Cognitive score is less able to identify certain problem areas (e.g. reading deficit). There were enough differences in the number of students identified by each

method that questions are raised as to which procedure is accurate.

Continuing to use aptitude scores, as directed by the test manual, to predict and identify learning disabilities would appear to be reasonable at this time.

Normative Comparisons

Except for a significant deficiency in Memory, students in this sample were average in all cognitive abilities and significantly below national norms in all achievement areas except Knowledge. Such findings are consistent with expectations for LD students. Case by case analysis reveals that many students had this pattern of average aptitude and below average achievement. However, at least one-third of this sample had scores with similar ranking in both Aptitude and Achievement. This suggests that the tests in both domains are examining similar skills. The results of the normative comparisons are in conflict with previous research findings that show LD students performing eight points lower on the average than the normative sample.

This is a time when efficient and effective diagnostic procedures are called for and needed. The WJPEB was neither efficient nor effective in identifying this sample of students with learning problems. Many of these students would not be receiving resource room assistance if the WJPEB had been the primary diagnostic tool.

Previous research suggested that the Broad Cognitive Ability score could be used to identify discrepancies, thus eliminating the time consuming calculation of Aptitude scores. This study reveals difficulties

associated with such calculations. The many inconsistencies in the case by case analysis and the reliabilities of difference score suggests that, for now, standard scoring procedures should be followed.

TABLE I
CASE BY CASE ANALYSIS

COUNT SOURCE	COUNT		
	Apt (%)	FSBCS (%)	MANUAL (%)
Moderate deficits in Reading	17	5	18
Severe deficits in Reading	14	14	13
Total deficits in Reading	31	19	31
Moderate deficits in Mathematics	12	10	12
Severe deficits in Mathematics	10	9	4
Total deficits in Math	22	17	16
Moderates in Written Language	0	5	12
Severe deficits in Written Language	0	1	2
Total deficits in Written Language	0	6	14
Moderate deficits in Knowledge	3	4	5
Severe deficits in Knowledge	4	5	0
Total deficits in Knowledge	7	9	5
Cases with 0 deficits	24 (34)	30 (42)	23 (32)
Cases with 1 deficit	34 (48)	32 (46)	31 (45)
Cases with 2 deficits	13 (18)	7 (10)	15 (21)
Cases with 3 deficits	0	1 (1)	2 (2)
Cases with 4 deficits	0	1 (1)	0
Total cases identified	47 (66)	41 (58)	48 (68)

TABLE II

ANALYSIS OF VARIANCE WITH REPEATED MEASURES:
WOODCOCK APTITUDE SCORES

<u>Source</u>	<u>df.</u>	<u>MS</u>	<u>F</u>	<u>P</u>
Bet Ss	65	723.69	7.76	.01
Within Ss	528	73.24		
Aptitudes	8	318.27	3.55	.01
Error	520	89.78		

TABLE III
CORRELATIONS BETWEEN ACHIEVEMENT SCORES AND
APTITUDE SCORES

SCORE	R Ach	M Ach	WL Ach	K Ach
FSBCA	.41**	.58**	.28	.62**
Rdg Apt	.44**	.43**	.39**	.61**
M Apt	.25	.49**	.25	.33**
WL Apt	.34**	.59**	.37**	.29
K Apt	.33**	.43**	.41**	.68**

**p<.01

TABLE IV
COMPARISON OF SAMPLE MEANS WITH NORMS

<u>Variable</u>	<u>Range</u>	<u>Mean</u>	<u>DS</u>	<u>Z</u>	<u>P</u>
FSBCA	63-134	97.04	13.12	-1.57	NS
Verb Apt	65-137	97.73	13.66	-1.27	NS
Reasoning	68-125	98.57	13.51	- .80	NS
V-P Spd	77-133	100.29	11.97	- .16	NS
Memory	69-121	95.43	11.89	-2.53	.05
R Apt	68-134	99.52	12.99	.27	NS
M Apt	79-127	99.90	11.14	- .06	NS
WL Apt	71-129	95.89	10.64	-2.31	.05
K Apt	64-129	96.40	11.33	-3.83	.01
R Ach	68-130	93.14	11.62	-3.85	.01
M Ach	66-132	95.12	12.09	-2.70	.01
WL Ach	78-125	92.40	9.90	-3.83	.01
K Ach	72-131	99.10	14.18	- .43	NS

REFERENCES

- Bolding, J. (1981). Statistics with finesse. Computer Program. Copyright Microsoft.
- Bracken, B., Prasse, D. & Breen, M. (1984). Concurrent validity of the Woodcock-Johnson Psycho-Educational Battery with regular and learning-disabled students. Journal of School Psychology, 22, 185-192.
- Breen, M.J. (1986). Cognitive patterns of learning disability subtypes as measured by the Woodcock-Johnson Psycho-Educational Battery. Journal of Learning Disabilities, 19, 86-90.
- Coleman, M.C. & Harmer, W.R. (1985). The WISC-R and Woodcock-Johnson Tests of Cognitive Ability: a comparative study. Psychology in the Schools, 22, 127-132.
- Compton, C. (1984). A guide to 75 tests for special education. California: Pitman Learning Inc.
- Cummings, J. & Moscato, E. (1984a). Research on the Woodcock-Johnson Psycho-Educational Battery: Implications for practice and future investigations. School Psychology Review, 13, 33-40.
- Cummings, J. & Moscato, E. (1984b). Reply to Thompson and Brassard. School Psychology Review, 13, 45-48.
- Estabrook, G. (1983). Test review. Journal of Psychoeducational Assessment, 1, 315-319.
- Green, B.F. (1981). A primer of testing. American Psychologist, 36, 1001-1011.
- Hall, R., Reeve, R., & Zakreski, J. (1984). Validity of the Woodcock-Johnson Tests of Achievement for learning-disabled students. Journal of School Psychology, 22, 193-200.
- Hessler, G. (1984). Use and Interpretation of the Woodcock-Johnson Psycho-Educational Battery. Allen, Texas: DLM Teaching Resources.
- McGue, M., Shinn, M., & Ysseldyke, J. (1982). Use of cluster scores on the Woodcock-Johnson Psycho-Educational Battery with learning disability students. Learning Disability Quarterly, 5, 274-287.

- Phelps, L. and Rosso, M. (1985). Validity Assessment of the Woodcock-Johnson Broad Cognitive Ability and Scholastic Aptitude Cluster Scores for behavior-disordered adolescents. Psychology in the Schools, 22, 398-403.
- Reeve, R.E., Hall, R.J., & Zakreski, R.S. The Woodcock-Johnson Tests of Cognitive Ability: Concurrent validity with the WISC-R. Learning Disabilities Quarterly, 1979, 2, 63-69.
- Shinn, M., Algozzine, B., Marston, D., & Ysseldyke, J. (1982). A theoretical analysis of the performance of learning disabled students on the Woodcock-Johnson Psycho-Educational Battery. Journal of Learning Disabilities, 15, 221-226.
- Thompson, P., & Brassard, M. (1984a). Validity of the Woodcock-Johnson Tests of Cognitive Ability: A comparison with the WISC-R in LD and normal elementary students. Journal of School Psychology, 22, 201-207.
- Thompson, P. & Brassard, M. (1984b). Cummings and Moscato soft on Woodcock-Johnson. School Psychology Review, 13, 41-44.
- U.S.O.E. (1977). Assistance to states for education of handicapped children: Procedures for evaluating specific learning disabilities. Federal Register, 42, 65082-65085.
- Wechsler, D. (1974). Manual for the Wechsler Intelligence Scale for Children-Revised. New York: Psychological Corporation.
- Woodcock, R. (1978). Woodcock-Johnson Psycho-Educational Battery. Boston: Teaching Resources.
- Woodcock, R. (1984a). A response to some questions raised about the Woodcock-Johnson I: the mean score discrepancy issue. School Psychology Review, 13, 342-354.
- Woodcock, R. (1984b). A response to some questions raised about the Woodcock-Johnson II: efficacy of the aptitude clusters. School Psychology Review, 13, 355-362.
- Woodcock, R. (1986). Assessment Service Bulletin #5. New York: Developmental Learning Materials.